

LESSON PLAN
INSTRUCTOR TRAINING COURSE

Title: Evaluation of student performance. Time required: 50 min.

Objective: To familiarize students with purpose and procedures involved in evaluating student performance.

References: 7-40.10/7 Evaluation of Student Performance.

Training Aid: Chart with four reasons for giving tests.

Presentation:

Motivation: An intelligence officer's life may depend upon his evaluation of student performance in the field. Examples can be an agent's mishandling of dynamite because he did not understand the importance of not letting it freeze, and an agent's faulty secret writing identifying the intelligence officer.

Explanation:

1. Reasons for giving tests.
 - a. Aid in improving instruction. *- prep. - instruct*
 - b. Provide incentive for learning. *- instruct to*
 - c. Provide basis for assigning marks. *- stand of prep. for marks*
 - d. Provide basis for selection and guidance. *- against need to measure ops. cap.*
2. Characteristics of a good test.
 - a. Valid -- closely related to job to be done.
 - b. Reliable -- give consistent results from time to time.
3. Types of tests.
 - a. Oral -- widely used in field, especially where there is one student for one instructor.

Document No. _____
No Change In Class. ☒
☐ Declassified
Class. Changed To: TS S C
Auth.: HR 70-2
Date: 5 OCT 1978 By: 72

OTL DONE COPY # 918

- b. Written -- probably has been over-used in many training situations.
- *c. Performance -- one of the most valuable techniques in testing indigenous personnel.
- *d. Observation -- should be systematized by means of check lists

Application: Allow students to bring up practical testing problems that have been or may be encountered in the field.

Examination: None.

Valid - measure what it is supposed to measure
Reliable - " accurately & consistently
Objective - several can get same score
- one interpretation by student
Discriminate - measure difference
Comprehensive - Sample all phases of instruction
Easily administered & scored -

7-40.10/7

EVALUATION OF STUDENT PERFORMANCE

General. Whether students are learning or merely attending classes, whether the instruction is effective or not, can only be determined by frequent and accurate tests. A good test program is essential to good training. All instructors are faced with the problem of constructing and giving tests to measure the progress of their students.

The basic concepts and techniques which the instructor will use in the examination stage of instruction are presented in this chapter. Discussion of the following points is included:

- a. The major purposes of testing.
- b. The forms of achievement tests and their uses.
- c. The characteristics of a good test.
- d. The construction of tests.
- e. The administration of tests.

Why give tests? The use of tests in training accomplishes these four major purposes:

- a. Aid in improving instruction. Properly constructed tests reveal gaps and misunderstandings in the student's learning. If frequent tests are given, such weaknesses can be discovered soon enough that the instructor can correct them by reteaching.

Tests aid in improving instruction in another respect. A test is actually a valuable teaching device in that students tend to remember longer and more vividly the points covered in an examination. Tests given periodically will encourage the students, as well as the instructor, to review the materials that have been presented and to organize the various phases of instruction into a meaningful set of skills, techniques, and knowledge.

Tests measure not only the student's performance, but also the instructor's. By studying the results of tests, the instructor can determine the relative effectiveness of his various methods and techniques. Weaknesses in his instruction will be revealed. It is particularly important that the instructor analyze test results and use these results in improving his instruction.

- b. Provide an incentive for learning. A basic principle of learning states that students learn more readily when they are held to account and made to feel responsible for learning. Generally, the instructor who gives frequent tests will find that his students will learn more. There is a danger, however, in playing up tests and test results as incentives to learning.

Revised Dec. 1956

-2-

Students who study primarily to pass tests usually forget what they learn much faster than those who are interested in earning because of the real values to be derived.

c. Provide a basis for assigning marks. Another purpose of testing is to determine which students have attained the minimum standard of performance and which have not. In many cases it is desirable to indicate the extent to which the student exceeds or falls before the standards required. Students learn difference amounts; the grade recorded for each student should be an accurate index of what he has learned. Unless a sound testing program is employed, it is impossible to determine the relative achievement of students.

d. Furnish a basis for selection and guidance. We are all familiar with tests which are especially designed to help in the initial classification of men and subsequent selection for various assignments. In addition to these tests, the results on training achievement tests furnish valuable supplementary information for selection and guidance of personnel. If tests are well constructed and actually measure student performance, the test results become a valuable basis for determining whether a student should be placed in a new job; whether he should receive advanced training; or whether he should be recommended for promotion to a job requiring greater ability.

Forms of achievement tests. There are four major forms of achievement tests. Each form has its specific uses, advantages, and limitations. In a well-rounded test program all four forms may be used. The forms of achievement tests are:

- a. Oral. Oral test questions enable the instructor to make an effective spot-check of his students' learning. They should be used in practically every lesson.
- b. Written. Written tests are of greatest value in measuring information. The short answer type of written tests have the added value of affording a rapid measure of students' knowledge over a large area of subject matter. Written tests can only indirectly measure a student's ability to apply knowledge and skills. In most training programs, other forms of achievement tests should be relied on to determine whether or not a student can actually perform a skill or technique.
- c. Performance. A performance test is a test designed to measure how well the student can perform a given piece of work. He is required to do something under controlled conditions while the speed and accuracy with which he performs the task are checked objectively. Merely asking a student to perform an operation and casually watching him is not a true performance test.

Performance tests are used to measure skill and information, ability to solve problems, and aptitude for training. An example of these, as used in Army service schools, is to require a student to wire and time an in-

-3-

line engine. It consists of directions to the student and to the checker, and a check list for recording the results of the student's performance.

A performance test is the best direct means of finding out whether a man can actually do a job and do it well. Written tests, no matter how carefully or cleverly constructed, may fail in this respect. For example, a student may pass a written test on how to reline brakes and fail in carrying out the actual task. Performance tests reveal specific difficulties better than any other type of test. A student is required to service, repair, shape, assemble, or disassemble something. A performance test is virtually the only way of revealing whether the student handles his tools efficiently, observes all necessary safety precautions, carries out the operations in the correct order, becomes emotionally upset when unable to do any part of the job, and fails to care for his tools properly when he has finished work.

Performance tests have certain disadvantages. First, they are difficult to set up properly. They are more difficult to administer than written tests since they often require that tools and special equipment be made available to the student. Much of the instructor's or assistant instructor's time is required in checking student performance. Whenever the performance is of a detailed or precise nature necessitating close observation, an assistant instructor is able to watch only one student at a time. In large classes, such performance tests are clearly out of the question.

In constructing performance tests, the instructor should examine the course of study and determine what portions of it should be tested by performance tests. He must provide the necessary tools and equipment; prepare directions for the students, stating the exact purpose of the test and operation to be performed; prepare directions for the examiner to follow while administering the test; prepare a check list for scoring the test; and, finally, try out the test on a few students or the other instructors and make the necessary changes.

SAMPLE PERFORMANCE TEST WIRE AND TIME AN ENGINE

PART I. (5 points) Time required to wire and time the engine.

Time Started _____ Time Completed _____ Total Time _____

SCHEDULE OF TIME AND POINTS EARNED

Within 15 min . . . 5 points	25 to 30 min . . . 2 points
15 to 20 min . . . 4 points	30 to 35 min . . . 1 point
20 to 25 min . . . 3 points	Over 35 min . . . 0 points

NOTE: Student to discontinue test if not completed in 35 minutes.

POINTS EARNED BY STUDENT (encircle one) 5 4 3 2 1 0

PART II. (4 points) ASSISTANT INSTRUCTOR: Check student on each of the following elements. Encircle the 1 (one) after the item in the YES column if the answer is yes; 0 (zero) in the NO Column if the answer is no.

CHECK WHILE STUDENT IS TAKING THE TEST YES NO

- | | | |
|--|---|---|
| 1. Battery ground cable disconnected until wiring completed | 1 | 0 |
| 2. Ground battery terminal tested for spark before connecting to battery clamp | 1 | 0 |
| 3. Engine starts the first time | 1 | 0 |
| 4. Student careful not to race engine | 1 | 0 |

POINTS EARNED IN PART I _____

POINTS EARNED IN PART II _____

TOTAL POINTS EARNED ____

Figure 1

d. Observation. Observation and observation techniques are of greatest importance in training because many phases of student achievement and behavior cannot be measured by the more formal kind of tests.

- (1) Observation tests are subject to several broad errors. The most common ones are:
 - (a) The error of lenience or severity: that is, rating all individuals too high or too low (frequently referred to as "personal bias" or "personal equation").
 - (b) The error of central tendency, or hesitation to give ratings at the extremes of a scale.
 - (c) The halo effect, or the tendency to rate a student either high or low in many traits because the instructor knows or thinks the student is high or low in some specific or particular trait.
 - (d) The so-called logical error which comes from presuppositions in the minds of the raters and lack of definiteness of the trait rated.
- (2) The effect of the above common errors can be reduced and judgments based on observation made more objective and valid if controlled methods of observing students are employed. The observer should:

- (a) Select phases of conduct that provide evidence of the quality being judged. If, for example, the problem is to evaluate students on leadership, the observers must see the students in situations that permit them to demonstrate their leadership ability, such as giving commands, directing activities of a small unit, and making and executing decisions. Similarly, to judge a student's ability as an instructor, he must be observed as he handles classes under varying conditions.
- (b) Make the observations comprehensive. He should avoid evaluating personnel on the basis of a few characteristics. All important phases must be considered, and there is only one way to do this; prepare a careful list of the important points.
- (c) Define the points to be observed. Each point must be clearly and accurately defined in terms of student behavior so that there will be no misunderstandings or ambiguities. Everyone concerned in the evaluation must think of the factors in the same way.
- (d) Define the standards of performance or conduct. It is not enough simply to look over a situation or watch students at work. The observer must know exactly what standards are to be expected as the student actually does the job. Work habits must be evaluated in terms of field use; personal characteristics must be evaluated in terms of present and future use. These standards should be reviewed by several instructors with varied backgrounds.
- (e) Observe accurately and impartially. An observer must be alert to all that is happening. He must avoid letting his opinions or biases influence his judgments; fatigue or emotional upsets should not enter the evaluation.
- (f) Make an accurate record immediately. Memory is too fleeting for an observer to trust in making important judgments. a complete record of the observation, notes on a check list, or ratings made at the time or immediately after the observation are necessary to make the judgment valid.
- (g) Combine judgments of several competent observers. Repeated observation increases the probability that all important factors will be considered.
- (h) Use check lists or observation check sheets. A check list for observation of performance, or a rating scale based on the consensus of competent judges insures that attention will be paid to the all-important points.

-6-

OBSERVATION CHECK SHEET

STUDENT _____ SECTION _____ DATE _____

ASSISTANT INSTRUCTOR _____

DIRECTIONS: Check the student Very Good, Satisfactory,
or Unsatisfactory.

ELEMENTS CHECKED VG SAT UNSAT

1. Work Habits
2. Speed
3. Accuracy
4. Care of Tools and Equipment
5. Observance of Safety Rules

TOTAL POINTS EARNED _____

ADDITIONAL COMMENTS: _____

Figure 2

Characteristics of a good test. An achievement test must have certain characteristics and meet certain requirements if it is to serve as an effective measuring device in a training program. In order to construct tests of high quality, the instructor must have a thorough understanding of these characteristics and requirements. There are six important factors which effect the quality of an examination. Although these are not considered to be separate and distinct factors, they are defined and discussed separately in order to develop a clear understanding of the features of an examination. Keep in mind that reliability, objectivity, and discrimination can be considered aspects of validity.

- a. The test must be valid. The test must measure what it is supposed to measure. This is the most important feature of an examination. A test designed to measure what the student has learned in a specific training program should measure his achievement in that training program and nothing else. If the test is constructed so that a highly intelligent student can figure out the answers whether or not he knows the subject matter involved, the test tends to measure general intelligence rather

-7-

than achievement. Other things, in addition to achievement, which often affect test scores are handwriting ability, reading ability, extent of vocabulary, knowledge of insignificant detailed information that has little to do with the subject matter of the course, the students' understanding of what the instructor is likely to ask on a test, and ability to memorize facts rather than the ability to apply and use them.

A practical method of estimating the validity of a test is to have several competent individuals carefully examine the content of the test to determine whether or not it includes materials taught in the course which are significant. The test results obtained should be compared with other measures of the students' achievement. One must keep in mind that a test may be highly valid for measuring certain outcomes. A variety of tests and other evaluating devices must be used in obtaining a valid measure of the various aspects of achievement.

b. The test must be reliable. A test is said to be reliable when it measures accurately and consistently what it does measure. If the test measures in exactly the same manner each time it is administered, and if the factors that affect the test scores affect them to the same extent every time the test is given, the test is said to be high in reliability. A highly reliable test, then, should yield essentially the same score when administered twice to the same student, provided, of course, that no learning occurs while the test is being taken the first time or no learning or forgetting place between testings.

There are several factors which affect the reliability of a test. In general, the reliability of a test can be raised by increasing its length. The more responses required of the student, the more reliable is the measurement of his achievement. The smaller the chances of guessing the correct answer to each item in the test, the greater the reliability.

c. The test must be objective. There are two aspects of objectivity which should be considered when constructing tests. The first is concerned with the scoring of the test. Several people should be able to score the test and get the same results. After the key has been made, personal opinion should not be a factor which affects the score.

It is obvious that the so-called old-type or essay tests, as usually constructed and scored, register poorly when measured by this standard. Instructors competent to judge rarely agree on the score that should be recorded for a given essay test paper. They do not have a common objective basis for marking. Instructors often notice that students are making higher marks than usual and begin to grade the test harder—that is, they begin to take off more points for errors. Such a system of scoring is not objective and should not be tolerated.

in a testing program. It is not to be assumed that the essay-type item should be eliminated altogether from tests. When carefully constructed, it measures well the ability of the student to organize and express his thoughts.

The second aspect of objectivity has to do with the student's interpretation of items in the test. Well-constructed test items lend themselves to only one interpretation by the students. Each student knows what he is required to do in responding to each item.

d. The test must discriminate. The test must be constructed in such a manner that it will detect or measure small differences in achievement or attainment. This is an absolute essential if the test is to be used for ranking students on the basis of achievement or for assigning marks.

Three things will be true of a test that meets this standard:

(1) There will be a wide range of scores when the test is administered to students who have actually achieved amounts that are significantly different. If the full range of achievement is to be measured, scores are likely to vary from the lowest to the highest possible score. However, for practical purposes in most subject matter fields, scores should vary from near the highest possible score to a score that is less than half of the total number of points on the test.

(2) The test will include items at all levels of difficulty. That is, the items will vary in difficulty from the most difficult one which will be answered correctly only by the best students, to an item so easy that almost all the students will answer it correctly.

(3) Each item will discriminate between students who are low and those who are high in achievement. Each item will be missed more frequently by poor students than by good students. If the good students are just as likely to miss an item as the poor students, the item does not measure in a positive direction. It has been found that poor students will answer correctly certain items that the best students miss rather consistently. An examination of these items usually reveals that they are ambiguous or are technically weak in other respects. The good students are able to detect the second, hidden, unintended meaning and allow it to lead them to make a response different from the one intended by the test maker.

As is true with validity, reliability, and objectivity, the discriminating power of a test is increased by concentrating on and improving each individual item in the test. After the test has been administered, a simple item analysis can be made which will readily show the relative difficulty of each item and, of greater

importance, the extent to which each discriminates between good and poor students.

e. The test must be comprehensive. It must sample liberally all phases of instruction which are supposed to be covered by the test. It is not necessary and it would not be practical to test every point that is taught in the course. A sufficient number of points should be included to provide a valid measure of the student's achievement.

f. The test must be readily administered and scored. Consideration must be given to the features of the test which make it readily administered and scored. It must be devised that a minimum amount of student time will be consumed in answering each item. The test items must also be constructed so that they can be scored quickly and efficiently.

Constructing a good test.

a. Selecting the best kind of measurement.

- (1) In selecting the best kind of measurement, the first thing an instructor must do is to decide on what to measure. In arriving at this decision, he must know what part of his instruction he expects his students to have mastered. The scope of a test should cover the attitudes, abilities, skills, principles, and facts presented in the lesson, job assignment subject, or course on which evaluation is to be made. These materials and skills should be stated in such a way that the students' performance will give an estimate of how well they can perform jobs in the field.
- (2) There are three steps that the instructor should take in deciding what to measure. The first is to analyze the objectives in terms of the general purpose of the training, making a precise indication in each case of how much the students are to get out of each objective. Next, he should list the various phases of the teaching assignment and indicate for each phase a definite formulation of the answers to these questions:
 - (a) What is the purpose of this phase of instruction?
 - (b) What are the students expected to learn in this phase of instruction?

Then, he should combine the first steps and outline just what phases are to be measured. All three steps are necessary in order to insure that all important points will

be covered by the test. The limited time allotted for testing requires the instructor to carefully select the tasks, questions, or problems he will use to check his students' mastery of the subject matter. The items should cover the important objectives and the essential skills and information. An examination samples the instructional materials but does not measure all that an individual knows about the subject matter. The value of the examination will depend upon how carefully the sample of the questions is chosen to cover the important objectives, skills, and information itemized in the instructional plan and job analysis.

- (3) After the problem of what to measure has been solved, the instructor should select the best kind of measurement for the specific purposes. For each testing purpose, some kinds of test or measurement are more effective than others. Knowing how to match the test with its purpose saves time for the instructor. If he has made a careful analysis of the teaching materials, if he knows what his objectives are, and if he knows some of the possibilities, and limitations of the testing techniques, he will have little difficulty in matching test methods and purposes. For example:
 - (a) If the instructor is to make certain that the students can make a wire splice in five minutes, he can use a performance test in which the student actually makes a wire splice.
 - (b) If the purpose is to see whether the students know how to work as members of a team in moving a command post, the instructor should observe the performance of a group of students engaged in that operation.
 - (c) If the purpose of the test is to determine whether the students know how to fill out a travel voucher, a written test can be used.
 - (d) If the purpose of the test is to determine the students' ability to express and develop their ideas on a tactical problem, the instructor can ask oral questions in class, have them submit a written report, or give them a short written essay assignment. The choice along these techniques would depend upon the importance and depth of the questions and the time available to both the students and the instructor.
 - (e) If the test is to measure the students' general information

about the subject matter of the entire course, a comprehensive selection of questions drawn from all the topics included in the course is needed. This can best be obtained by using a test consisting of a variety of questions requiring little or no writing on the part of the students.

- (4) There are some practical limitations to be considered in selecting the best type of measurement. An evaluational or measurement situation should be selected in terms of the exactness with which it fits the specific testing purpose. However, the demands of the training situation may require a compromise between the type of test best fitted for the job and the one that can be used. For example, a shortage of time or equipment may not permit the use of a performance test. Furthermore, an instructor's experience in constructing examinations and in selecting the important points to observe in a student's performance will influence the type of test he will use. Finally, the type of personnel being trained must be considered; for example, men with little schooling will, in general, not be at their best on a paper and pencil test.
- b. Organization of the test. Whether the instructor is planning to use a series of oral questions, a written test, a performance test, or a combination of these, he should organize the testing materials to the best advantage.
- (1) The wording of the questions, the directions, and the instructions for doing the job must be simple, clear, and concise.
 - (2) The test should be reasonable.
 - (a) If the test is to measure speed of performance, the time limit should be short enough to allow only the fastest workers to finish.
 - (b) If everyone is expected to attempt all the items, the number and the type of operations or problems should be selected with this purpose in view.
 - (c) The test should deal with those materials, attitudes, or operations which have been taught or which the students have had an opportunity to acquire.
 - (3) A limited number of different item types should be used. Not more than four or five kinds of test item—oral, performance, essay, multiple choice—should be used on one test. Directions for each kind must be included.

- (4) The answer spaces for written items of the objective form should be arranged along one side of the page, or in such a way that the scoring can be handled by means of cut-out stencil or scoring sheet; or separate answer sheets should be provided.
 - (5) A test measuring mastery of lesson fundamentals should include only items relating directly to the essential attitudes, skills, or information of the lesson that the men will need on the job for which they are being trained.
 - (6) A test measuring degree of mastery of a subject or course should include a few easy items as well as a few very difficult ones so that no student will make a perfect score, nor any make a zero score.
- c. Content check. Before the test is used, it should be submitted to instructors familiar with the subject matter for review and criticism. This procedure is necessary in order to insure the elimination of undesirable emphasis upon unimportant phases of the instructional materials. The reviewing instructors should evaluate the test items, directions, and organization in terms of the
- (1) purpose of test.
 - (2) characteristics of a good test.
 - (3) proper emphasis on "must know" and "should know" instructional materials.
 - (4) type of personnel being trained.
 - (5) unity, coherence, and proper emphasis in the wording of questions or directives.
- d. Time check. In all tests, especially written ones, a competent instructor or other qualified individual should actually take the test in order to determine the time required for administering it. Students should not be expected to complete a test in as short a time as an instructor; on the other hand, a test may be found to be too short.

Administration of tests.

- a. Need for careful administration. A poorly administered test is worse than no test. The results of improperly administered, as well as poorly constructed, tests give the instructor an incorrect impression of his students' mastery of the materials taught. The results of

this misconception may not be apparent until it is too late. The only way to prevent this error is to check students frequently with carefully constructed tests properly administered and interpreted.

In order that a test may measure mastery of the materials taught, the students must be given every opportunity to do their best. If a man gets a low score because he misunderstood the instructions or was fatigued or emotionally upset, his score is not a true indication of his ability. Likewise, a high score resulting from cheating or from the improper help of the instructor or his assistants will be a false indication of his ability.

b. Procedure in administering tests. The instructor who is giving any test should:

- (1) Have all testing materials ready. An instructor who has prepared the items for the examination or the problems to be solved must be sure that enough test blanks, directions, check lists, operation sheets, tools, pencils, scratch paper, or any other material required will be on hand in the classroom at least 15 minutes before the test is scheduled to start. The testing materials may be distributed to the men either as they enter the room or when they are ready to start. The instructor may delegate the duty to one of his students, but he is responsible for seeing that the materials are ready for use.
- (2) Train any assistants needed. For the proper administration of performance tests, for those requiring considerable time, or for any test being given to a large group of students, the instructor should be provided with assistants.
 - (a) There should be one assistant for every 25 men tested on a written examination which covers materials of a job assignment, subject, or course. A performance test will usually require an assistant for every one or two men.
 - (b) The smoothness of testing procedure will depend upon the efficiency of the assistants. The instructor should work out a convenient schedule for the assistants to follow in distributing and collecting test materials, in seating and dismissing the students, in checking their work, supervising the conduct of the test, and in giving them any help that is proper.
 - (c) The instructor should go over the examination carefully with assistants and indicate to them the points at which he may expect the students to have difficulty, the amount and kind of help they may give the students, and their exact function and position in the classroom.

- (3) Provide the best possible testing conditions. Students are not able to do their best in a dark, noisy, or crowded classrooms. The instructor should eliminate all interest-deströying factors, arrange the mechanical aspects of the room in the most expedient way, and place the seats (or working layouts) so that each man will have ample working space and be far enough from his neighbor that there is not temptation to see the work of anyone else. Men should be mentally and physically rested before they are given any crucial examination. No one in a state of fatigue, such as might be brought on by a long march or a sleepless night, can do justice to an examination.
- (4) Give students a good start. A test, like any other phase of instruction, should be started in a businesslike manner. The instructor is responsible, first of all, for putting the men at their ease. In his directions explaining the test situation, he should encourage the men to do their best.
 - (a) The instructor must make certain that the men understand the test directions. General instructions for test of any length should be written out. These directions should be read clearly and unhurriedly to the men who should be encouraged to ask questions at the end of the reading in order to clear up any possible misunderstandings.
 - (b) Before starting the test, the instructor should indicate to the students the amount of kind of help they can secure and the materials that are to be used. They should be told whether a single over-all time limit, time limits on separate items or sections, or credit for speed of work will be used. If a test is being administered with no time limit, the students should be told that they will have as much time as they need.
- (5) Conduct the test properly. In order that a test may give the best possible indication of the ability of each man in the class, the conditions under which the test is conducted should affect each man alike.
 - (a) Order in the classroom and on the job must be maintained. No student should be allowed to interrupt another student or create any disturbance.
 - (b) Some tests are timed. They must be timed accurately. For most purposes this is best done by announcing in advance the exact time at which the men will cease work on the test.
- (6) Conduct a discussion and review of every test. The teaching value of a test is not fully realized unless a discussion and review

-15-

of the students' performance on the test is conducted. Such discussions give the instructor an opportunity to clear up any misunderstandings and to fill in any gaps in learning.

A POORLY PLANNED TEST

1. I'd kind of like to know how much you fellows got on the M1 rifle you studied last week. Let's see. I guess I can pick out a couple of things for you to do to show if you were on the ball. Let me see.....
2. "Ah yes! Let's use these two guns up front. Suppose you come up two at a time and disassemble the trigger and interlock mechanism on these guns and then you can put them back again."
3. "Don't worry about how you'll be marked just do the job and you'll get what you deserve. If you're on your toes, you'll come out all right. If you don't know a trigger from the rear end of a mule, you'll get quite a kick out of your grade. I'll be up here watching everyone of you."
4. "I think we can finish this test in an hour if you don't drag your tails. Anyway that's all the time we have. I guess these two guns are enough. Say, maybe we'd better put these guns on the tables in the rear so you men not working on the test can do some work on this model up front. I want the men not working on the test to keep working until they are called. Johnson, lend a hand here. And I guess maybe you'd better help me during the test. I'm not sure how this thing will work."
5. "Now you fellows all know what to do. And Johnson, suppose you mark the men who work on this second gun. Just use your model on this thing and I'll help you as we go along. Come on fellows, we've lost fifteen minutes already and we've got to finish today. Let's get going."

ANALYSIS OF A POOR TEST ON THE M1 RIFLE

GETS AN INSTRUCTOR INTO DIFFICULTIES

1. With the class in front of you waiting for a test, it's kind of late to start thinking about the skills you want to check. Well, maybe you'll think fast and hit on something important anyway.
2. Just two little jobs to test the most significant skills on this gun? Hardly enough! You can't tell much about a barrel of apples by testing just two little ones off the top.
3. Might as well get the grades right out of the dream book. Measuring performance on this basis is like measuring a yard with a rubber band. What counts in the score here? Speed? Accuracy? Identification of parts? I.e., knowing a trigger from the rear end of a mule? If anyone knows they're keeping it a secret.
4. Heading for the rocks. No clear idea as to whether the equipment is adequate to do the job in the time available. Time lost and confusion present in setting up equipment, all of which should have been arranged at fixed stations in advance of the test period. Only a guess on the number of assistants needed. You wouldn't organize a TD so haphazardly. Your strategy (knowing what your objective(s) is) and your tactics (knowing how to achieve that aim) must be carefully planned in advance.
5. The students know they have two tests to do, but they don't have a cloudy idea as to what will be looked for when they do them. Shall they work fast and risk errors? Slowly and accurately? As for the assistant, the ceiling is zero with more bad weather ahead.

BECAUSE THESE QUESTIONS WERE NOT RAISED AND ANSWERED

1. What major skills were to be developed by studying this gun? Repair of the gun? The firing of the gun? Assembly or disassembly? This is the first problem to be solved.
2. What are the most significant tasks related to these skills which indicate whether the necessary proficiency was developed? Choose enough representative tasks to really test the men.
3. What will be the standards by which you will judge good or poor performance? In scoring, will you judge speed of performance? The quality of the end product? Some other factors? Are these features the important aspects of the performance which will govern good or poor work when the men go on the job.
4. How can the administration of the test be organized most effectively? Make a shop layout plan and try the test ahead of time to determine (a) how much time it will take to give the test, (b) what equipment will be needed, (c) where equipment should be stationed, and (d) how many assistants will be needed. The trial of the test should also help you check on the effectiveness of the standard you have set for scoring. Organize the test so every man is busy for the whole period.
5. Do students, timekeepers, scorers know exactly what to do and how to do it? Do the students know on what basis they will be scored? Do the scorers know how to score? Do they have standard score sheets? Will everyone be scored in the same way on the same task?